2017-09

Comparing two tools for mobile-device forensics

Martin, Casandra M.

Monterey, California: Naval Postgraduate School

http://hdl.handle.net/10945/56153

NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

COMPARING TWO TOOLS FOR MOBILE-DEVICE FORENSICS

by

Casandra M. Martin

September 2017

Thesis Advisor:	Neil C. Rowe
Second Reader:	Michael McCarrin

Approved for public release. Distribution is unlimited.

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704–0188

1. AGENCY USE ONLY *(Leave Blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	September 2017	Master's Thesis 09-20-2015 to 12-19-2016

4. TITLE AND SUBTITLE

COMPARING TWO TOOLS FOR MOBILE-DEVICE FORENSICS

5. FUNDING NUMBERS

6. AUTHOR(S)

Casandra M. Martin

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Naval Postgraduate School
Monterey, CA 93943

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

N/A

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

The views expressed in this document are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. IRB Protocol Number: N/A.

12a. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release. Distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(maximum 200 words)*

Gathering forensic data from mobile devices has become essential with the rise of mobile technology and the value of the data they store. This thesis looked at a new analysis platform, which we called "T," and compared its output with an existing tool, Cellebrite's Physical Analyzer (CPA). We imaged 22 different devices with Cellebrite's imaging software and then analyzed the images with both tools. The phones were categorized into 1 of 7 categories based on their content and usage. We concluded that CPA and T have different benefits. CPA was strongest in its user interface and ability to determine web usage, as well as being able to analyze a variety of devices. T had the ability to allow for keyword searches, which allowed us to be able to identify more email address possibilities. We propose testing more recent updates of the tools against a larger corpus of phones in future work.

14. SUBJECT TERMS

Mobile Forensics, Mobile Phone Imaging, Mobile Analysis Tool, Android, iOS, BlackBerry, Cellebrite, Email, Web, UFED Touch, and OS

15. NUMBER OF PAGES 47

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UU

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2–89)
Prescribed by ANSI Std. 239–18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release. Distribution is unlimited.

COMPARING TWO TOOLS FOR MOBILE-DEVICE FORENSICS

Casandra M. Martin
Civilian, Department of the Navy
B.S., California State University, Monterey Bay, 2014

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN COMPUTER SCIENCE

from the

NAVAL POSTGRADUATE SCHOOL
September 2017

Approved by: Neil C. Rowe
Thesis Advisor

Michael McCarrin
Second Reader

Peter J. Denning
Chair, Department of Computer Science

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

Gathering forensic data from mobile devices has become essential with the rise of mobile technology and the value of the data they store. This thesis looked at a new analysis platform, which we called "T," and compared its output with an existing tool, Cellebrite's Physical Analyzer (CPA). We imaged 22 different devices with Cellebrite's imaging software and then analyzed the images with both tools. The phones were categorized into 1 of 7 categories based on their content and usage. We concluded that CPA and T have different benefits. CPA was strongest in its user interface and ability to determine web usage, as well as being able to analyze a variety of devices. T had the ability to allow for keyword searches, which allowed us to be able to identify more email address possibilities. We propose testing more recent updates of the tools against a larger corpus of phones in future work.

THIS PAGE INTENTIONALLY LEFT BLANK

Table of Contents

List of Tables

THIS PAGE INTENTIONALLY LEFT BLANK

List of Acronyms and Abbreviations

API Application Program Interface

App Application

CPA Cellebrite's Physical Analyzer

DFU Device Firmware Update

DFXML Digital Forensics Extensible Markup Language

DoD Department of Defense

FOUO For Official Use Only

FTK Forensik Toolkit

GUI Graphical User Interface

iOS iPhone Operating System

MDS Mobile Data Service

MIDP Mobile Information Device Profile

NIST National Institute of Standards and Technology

NPS Naval Postgraduate School

OS Operating System

SIM Subscriber Identity Module

SMS Short Message Service

UFED Universal Forensics Extraction Device

URL Uniform Resource Locator

USB	Universal Serial Bus
WAP	Wireless Application Protocol
XML	Extensible Markup Language

Acknowledgments

First and foremost, I would like to thank my thesis advisor, Dr. Neil Rowe. My thesis would not have been possible without your help and guidance. Thanks to my second reader, Michael McCarrin, for not only being my second reader but also for being very present and available whenever I needed help in any subject. Also, thank you specifically to two of my classmates, Jennifer Johnson and Johanna An: you were there alongside me every step of the way being my support system when I needed it most. To my friends and family: thank you for putting up with me during the ups and downs of this process. Mom and Dad, my degree is dedicated to you. I could not have done it without your love and support and for that I will be forever grateful.

There were some very influential people throughout my academic journey who I would also like to thank. Joe Welch was my first Computer Science professor at Hartnell Community College. Without his persistent efforts to convince me to take my first computer science class, I would have never got to this point. Alison Kerr was my first introduction to NPS and has taught me so much. Because of her I gained and honed skills I never knew I could have or need. She was always there to give me advice about everything and anything under the sun. Thanks also to Sue Higgins for being a great role-model and bringing Cyber Adventurers into my life, Warren Yu for always lending an ear and being a fountain of wisdom, Riqui Schwamm for always answering any questions and e-mails and for all his help throughout my thesis, and to all my professors for two great years of lectures, homework, tests, and sleepless nights.

I would also like to thank the Scholarship for Service Program funded by the National Science Foundation who allowed me to attend NPS full time as a civilian student. Thank you to Cynthia Irvine and Mark Gondree and all the people who worked behind the scenes to make my time at NPS possible.

To everyone who supported me and was there for me throughout my time at NPS, thank you from the bottom of my heart.

THIS PAGE INTENTIONALLY LEFT BLANK

CHAPTER 1:
Introduction

Forensic analysis of files and systems is a useful way of characterizing large volumes of digital data. With the rise of mobile technology and the amount of data mobile devices now hold, it is important to be able to analyze the digital data within these devices. Additionally, deriving metadata from bulk mobile data has become increasingly beneficial since a vast majority of communications now occur via mobile devices. Digital forensics tools, such as Cellebrite, are necessary to be able to extract and analyze data content. These tools have served their purpose well and have improved over time.

This thesis will look at a fairly new digital forensics analysis platform, which we refer to by the alias "T." It will discuss the differences and similarities in T's capabilities for mobile phone image analysis with the capabilities offered by Cellebrite's Physical Analyzer.

We will image a variety of mobile devices that have been collected from many different countries and attempt to gather specific data from them.

1.1 Contribution to Department of Defense

This research will provide an understanding of the T tool and its capabilities in regards to accurately analyzing data found on mobile phones, specifically iOS and Android devices. It is crucial to be able to quickly and effectively analyze mobile devices that may contain information related to national security. Preferably, we would do this using open source tools.

1.2 Scope

The scope of this thesis will be limited to a comparison of information that can be obtained from mobile images using T's mobile analysis tools with the information that can be obtained using Cellebrite's Physical Analyzer Software. We will provide an analysis of the T tool and its performance in comparison to Cellebrite's.

1.3 Research Questions

Through this thesis, we aim to answer the following research questions:

1. Are there identifiable differences between Cellebrite and T with respect to mobile device analysis capabilities?
2. Can we gather data from these files using T's mobile device image analysis tool?
3. Can the same be done for files on an Android device?
4. Are there files found by one tool that are not found by the other?
5. Are there email addresses found by one tool and not the other?

1.4 Thesis Structure

The remainder of this thesis is organized as follows. Chapter 2 will discuss some background information on mobile forensics tools and related work on this topic. Chapter 3 will cover the methodology and experimental process. Chapter 4 will discuss the experimental results and findings. Chapter 5 will end with conclusions and future work.

CHAPTER 2:
Background and Related Work

2.1 Mobile Device Use and Evolution

Nearly two-thirds of Americans are now smartphone owners as of April 2015, which is a 35% increase from 2011 [1]. At the same time that consumers have been increasing their purchase of and use of mobile devices, manufacturers have been increasing the storage capacities of these devices. This permits users to store more data and information than ever before [2]. Mobile devices are essential these days for the average American: they are used to communicate and provide instant information wherever you are. Eighty percent of mobile device users report using their devices to access the Internet and download content [3]. With all this use of mobile devices to communicate and facilitate our lives, it is no wonder that they are rich in personal and valuable information.

2.2 Mobile Forensics

"Mobile forensics is a branch of computer forensics that focuses on mobile devices, typically smart phones, tablets, iPads, and cellular devices" [4]. It is a type of electronic data gathering, which targets taped conversations, pictures, texts, emails, phone numbers, video, etc. [2]. Just as computer information is hard to delete, since data can only be truly deleted by overwriting, the same applies to mobile devices. Users may believe data is permanently gone once deleted, but often is recoverable and reviewable by forensic examiners [2], [4].

2.3 Guidelines

Mobile forensics is a fairly new and growing subarea of computer forensics, so the tools and resources are in the early stages of maturity [5]. The National Institute of Standards and Technology (NIST) provides a guideline that discusses procedures for the preservation, aquisition, examination, analysis, and reporting of digital evidence [6]. This is not meant to be a step-by-step guide on how to perform forensic examination on a mobile device, but rather it is meant to be a starting point and to outline the important principles of mobile forensic examination. The guide is meant to be used by law enforcement, incident responders, and other types of investigators. It addresses common circumstances that may be encountered by organizational security staff [6]. NIST Special Publications tend to be a good source and starting point on computing topics because they are generally accepted as the baseline standard.

2.4 Mobile Operating Systems

"A mobile operating system is an operating system that is specifically designed to run on mobile devices" [7]. On a desktop or laptop, an operating system like Linux or Windows is responsible for making physical resources (such as RAM, secondary storage, displays, etc.) available to the system software. Similarly, "a mobile operating system is the software platform on top of which other programs can run on mobile devices" [7]. There are many different types of mobile operating systems and they are constantly changing, which means an operating system that is available now most likely will not be available after a few years [8]. Since compatibility with a forensic tool is based on the mobile device's operating system and there are so many, each with multiple versions, determining compatibility can be a challenge [9]. Three of the more common mobile operating systems are briefly described below.

2.4.1 Android

The Android operating system is developed by Google, and it was originally released in September of 2008. "It is based on the Linux Kernel and is designed primarily for touchscreen devices such as smartphones and tablets. Android has the largest installed base of all operating systems and has been the best-selling mobile operating system since 2013" [10]. The source code is open-source and is developed in private by Google and then released publicly when a new version comes out [10]. "The Linux Kernel provides access to core services such as security, memory management, process management, network stack, and driver model. Because it is open-source it is designed to simplify the reuse of components since developers are given full access to the same framework APIs used by core applications" [9]. The use of a Linux Kernel in Android phones provides an advantage because there is an ability to use Linux commands such as "dd" when the mobile device is rooted. The downside to this is that the security features make forensic analysis more difficult [11].

2.4.2 iPhone

"The iPhone runs an operating system called iOS. It is a variant of the Darwin operating system that is also found in Mac OS X. The operating system takes up less than half a gigabyte" [12]. It only supports applications distributed through Apple's App Store. The operating system is managed and updated through a system known as iTunes from a computer. Apple provides free updates through this system as long as the required version is being used [12]. "The iPhone operating system has four layers; the core OS, core services, media, and Cocoa Touch. The core OS and core services are the bottom two layers and they contain the fundamental interfaces for iOS. These include the interfaces for accessing files, low-level data types, network sockets, and the UNIX sockets" [9].

2.4.3 BlackBerry

"The BlackBerry OS is a proprietary mobile operating system developed by BlackBerry Limited. The operating system provides multitasking and supports specialized input devices that have been adopted by BlackBerry. The platform is best known for its native support for corporate email through MIDP 1.0 and 2.0 which allows synchronization with Microsoft Exchange, Lotus Domino, and Novell GroupWise email" [13]. The operating system supports WAP 1.2 and it gets updated automatically whenever it has access to a wireless Internet connection [13]. There is little public information known about the BlackBerry operating system architecture. What is known is that it is run on a VM or virtual machine with Java. Proprietary and MDS are the two runtime environments the BlackBerry operating system has [13].

2.5 Other Mobile Forensics Work

There was a similar project done by the University of Glasgow where a group of researchers collected re-sold mobile devices and attempted to gather data from them [14]. They looked at two aspects; the first was how much sensitive information they were able to gather from these devices and the second was the consistency of the information gathered from different forensic applications [14]. They found that the smartphones contained some sensitive data, but not as much as they expected, and of the three software products tested, two performed significantly better, producing similar results [15].

2.6 Previous Tools

Since mobile devices are constantly changing there has been difficulty with digital forensics tools being able to keep up. Some popular tools are:

1. **FTK Mobile Phone Examiner.** This tool was the most commonly used forensics tool in the U.S. in 2011. Data could be collected off a mobile phone via cable, Infrared, or Bluetooth without modifying any content on the phone [16].
2. **Oxygen Forensic Suite.** This tool is Europe's preferred mobile forensic tool. It has all the abilities that many other tools have, but additionally it could provide geo-tagging information for Nokia phones. Not many other tools could do that, so that makes them stand out [17].
3. **EnCase Neutrino.** This tool was similar to the Cellebrite tool we used because it also allowed for a connection via USB where the tool identified the device and provided all possible adapters. This tool imaged the SIM cards, providing user-account data as well [16].
4. **Paraben's Device Seizure.** This tool was special in that it had low system requirements. It was able to run on any computer no matter if it was old or new [17].
5. **iPhone Analyzer.** This tool supports iPhone 5 and older. It uses Apple's own iTunes software to download the Analyzer via the iTunes App Store and is able to recover backups, geo-locate the device, view all photos, examine the address book, and export files to a local file system [18].

2.7 Mobile Triaging

Triaging in medicine means deciding when patients get seen based on the urgency of their condition. As a general definition, triage is the process through which things are ranked in terms of importance or priority [19]. With the increasing popularity of mobile devices and many malicious people using them for crimes, there is a strong demand for efficiently accessing the data of value on mobile devices [20].

Before, mobile analysis consisted of manual inspection and pictures taken of phone screens, but that has completely changed due to the fast pace of mobile technology and the forensic tools that are now available. To figure out what devices are worth looking at and which will not be too helpful, analysists need a way of distinguishing them. This is where automatic triaging and categorization comes into play [20]. Work on data mining and machine learning has helped advance the ability to triage mobile devices and more efficiently find the content that would be of value on mobile devices [20].

Machine learning and data mining algorithms have played a major role in mobile triaging. A collection of known and categorized phones serve as a training corpus to then be able to classify new phones based on features and phone content [21]. There is a technique called "5 minute forensics" that has served as a framework for mobile triaging. This technique uses five pre-determined categories that refer to amount of usage ranging from occasional to hacker [21]. The idea is that if one device gets classified as "occasional," meaning little to no usage, and another as "hacker," meaning a large amount of usage, then the obvious one to look at first is the latter one because it was used more and might contain more data of value.

THIS PAGE INTENTIONALLY LEFT BLANK

CHAPTER 3: Methodology

In this chapter, we provide more details about the Cellebrite Physical Analyzer tool and the T mobile analysis tool and the approach taken to evaluate them. We will describe the experimentation process, failures, and successes.

3.1 Device Imaging

To do any analysis on a mobile device, aside from physical inspection of the device, it is necessary to create an image of that device. An image is a copy of the contents of the device that is transferred to another device such as a computer or laptop.

3.1.1 Data Acquisition Techniques

There are two main approaches to doing a mobile extraction, physical and logical. A physical extraction is a bit by bit copy of memory. It includes flash memory which allows access to data and files that might have been lost or deleted. [22]. A logical extraction is not a bit by bit copy; it is more of a data request. The device's own API is used to communicate with it and data that is live and viewable on the device can be requested. The device then replies and sends the data over a communications channel. A logical extraction is much quicker since there is a lot less data to gather [23]. There were a few devices that did not allow for a physical extraction, so for those devices we decided to do a logical extraction. For this thesis we mainly performed physical extractions.

3.1.2 Cellebrite UFED Touch

For this thesis, we used Cellebrite's Universal Forensics Extraction Device Touch hardware [24]. The UFED allowed for several different mobile device types to be attached and imaged. The hardware worked alongside Cellebrite's Physical Analyzer Software which needed to be run simultaneously to image the device. In our data set there were many different devices that required many different attachments to be able to access them. The UFED came with all possible attachment options.

Once the right attachment was found the device needed to be fully charged before imaging could be attempted. The UFED provided a set of specific instructions to prepare each type of device for imaging. We focused on mobile phones that allowed for a physical extraction.

The physical extraction process varied from phone to phone. Generally, the imaging process, with the exception of iPhones, was as follows:

1. Enable debugging; this was done manually if necessary.
2. Turn off the phone and plug it in to the UFED hardware via a USB connection.
3. Plug the UFED into the USB port of a computer or laptop running the Cellebrite Physical Analyzer Software.
4. Follow the prompt provided by the UFED to start the extraction process via the software running on the computer.

After these steps were taken the extraction process began and extracted a bit-by-bit memory copy to a file path of choice.

The imaging process for an iPhone device was different than the process for other phones. All iPhones had the same set of instructions. The process for iPhones typically went as follows:

1. Turn off the iPhone.
2. Put the iPhone into DFU mode according to instructions on the screen.
    a. Hold the Home button and plug the iPhone in via a USB cable.
    b. Keep holding the home and additionally the power button down at the same time when an iTunes image appears on the screen.
    c. Keep holding both buttons for 3 seconds after the screen goes black.
    d. Release the power button. At this point the iPhone has entered DFU mode.
3. Observe the iPhone's information that appears on the screen. Notice that the serial number, OS version, and whether or not it has been jailbroken appears onscreen.
4. Continue the extraction process and select the Physical Extraction option.
5. Select the file path where the extraction should be placed.

The imaging process for BlackBerry phones was similar to the Android imaging process except the phone did not need to be turned off. The rest of the steps were the same. The

Blackberry phones imaged much more quickly than most of the Android phones.

3.2 Mobile Image Analysis Tools

After the device was imaged and the extraction process was complete, the image needed to be analyzed. This was done with mobile image analysis tools. Our goal was to evaluate the effectiveness of T. To accomplish this we compared the analysis of a device using Cellebrite to the analysis of that same device using T. Specifically, we were looking for differences in email addresses and web usage data between both analyses.

3.2.1 Cellebrite Physical Analyzer

Cellebrite's UFED Touch came paired with Cellebrite's Physical Analyzer [25]. The software was used to both extract the data from the devices as well as view the content once the extraction was complete. Its GUI was user-friendly and provided a filesystem type of view with files and folders off to the left hand side. The various types of files such as pictures, emails, media, contacts, accounts, etc. were listed and it provided the number of each found. Clicking on the file type opened a tab listing all the files and information on all those files.

Cellebrite provides an option to create a report for any imaged device. The report can include all files found on a device along with hash functions computed on files. This report can be exported in various formats. We chose to export the reports in XML format.

The Physical Analyzer produces reports in a proprietary XML format. We converted these XML reports to DFXML to enable use as input to other scripts and tools that run analysis on the mobile device images. Conversion was performed using an existing Python script that was written by Riqui Schwamm and Dr. Neil C. Rowe from NPS. "DFXML stands for Digital Forensics XML and is an XML language designed to represent a wide range of forensic information and forensic processing results" [26]. DFXML is a standard that comes from The National Institute of Standards and Technology (NIST). NIST uses DFXML internally for some research projects and to distribute some information [27].

3.2.2 T

T is the alias we have assigned to a mobile forensics tool that has been classified as For Official Use Only or FOUO. T is basically a version of Autopsy with a few additional features. "Autopsy is a digital forensics platform and graphical interface to The Sleuth Kit and other digital forensics tools. It is used by law enforcement, military, and corporate examiners to investigate what happened on a computer or device" [28]. The T interface is GUI based. It is similar to Cellebrite's in that it is set up like a file system. The additional features include some extra modules, including the Bulk Extractor module, Smirk module, Volatility module, and Forensic Toolbox module. For our experimentation we used all of these modules.

T allows a user to add data sources to a case as input. For our data sources we added either the binaries or disk images extracted using the UFED touch. There is no limit to the number of sources that can be added to each case. We created a case for each mobile device.

3.3 Phone Corpus

Our data set consisted of 20 mobile phones and 1 Apple device (iPod) that came from the Real Data Corpus, all imaged using Cellebrite's UFED Touch. Five of those mobile phones were iPhones, 5 were Samsung, 2 were BlackBerrys, 1 was HTC, 2 were LG, 1 was Motorola, 3 were Nokia, and 1 was Sony. Table 3.1 shows the details on the phones that were imaged. The first two letters of the phone names are the country code of the phones country of origin.

Table 3.1. Phone Corpus Details

Phone	Vendor	Name	Model	Extraction Type	OS	Version
BZ-12	Samsung	Galaxy S III	GT-I9305	Physical	Android	4.1.2
BZ-25	Samsung	Galaxy Ace 3	GT-S7270L	Physical	Android	4.2.2
CA-01	Apple	iPhone	4	Physical	iOS	5.1.1
DE-18	Motorola	Razor	GSM V3	Physical	Android	2.3.6
FR-04	Nokia	Lumnia	1520	Logical	Windows	8
FR-05	Apple	iPhone	4	Physical	iOS	4.3.2
IN-11	Dell	ZTE Blade	XCD35	Physical	Android	2.2
SG-27	Samsung	Galaxy III	GT-I5801	Physical	Android	2.1
SG-28	LG	Pop	GD510	Logical	Flash	n/a
SG-29	Nokia	N97 mini	N97 mini	Physical	Symbian	9.4
SG-34	Samsung	Corby Pro	GT-B5310r	Logical	Proprietary	n/a
SG-50	HTC	Incredible S	S710e	Physical	Android	2.2.1
SG-64	LG	Optimus L3	E400	Physical	Android	2.3.6
SG-66	Nokia	X3	X3	Physical	unknown	unknown
SG-80	Apple	iPhone	2	Physical	iOS	3.1.3
SG-81	Apple	iPhone	3	Physical	iOS	5.1.1
SG-88	Apple	iPod	3G	Physical	iOS	4.2.1
TH-02	Sony	Xperia	E15i	Physical	Android	2.1
TH-05	BlackBerry	Curve	9300	Physical	BlackBerry	5.0.0.912
TH-09	Samsung	Ch@t 322	GT-C3222	Physical	Android	n/a
TH-12	Apple	iPhone	3G	Physical	iOS	4.2.1
TH-20	BlackBerry	Curve	9300	Physical	BlackBerry	6.0.0.546

Here, we list the specifications of all the imaged devices including whether they had a physical or logical extraction.

3.4 Mobile Image Inspection and Content

All device images were analyzed using Cellebrite's Physical Analyzer as well as T. We compared and contrasted the outputs of each tool. We focused on email and web usage. We used the information gathered on these files as our basis for determining the strengths and weaknesses of the two tools. Web and email files are common in most devices and provided a good baseline. Real email addresses have been replaced with equivalent addresses for privacy reasons.

3.4.1 Analysis using Cellebrite

With the Cellebrite's Physical Analyzer Software the process of gathering email addresses varied. On some devices the tool did a good job collecting them and gathering them under the email tab. It allowed us to navigate the addresses found and then showed us where on the device they were found.

There were devices that provided zero addresses in the list of emails. Deeper inspection and searching through the logs and files showed that there were indeed some email addresses present.

Facebook Messenger seemed to provide email addresses on most devices that contained Messenger data. Account data and email were recorded among the message exchanges between the user and other contacts.

CPA was able to provide the device logs, which recorded all activity on a device and were a good resource when the tool had not been able to find much information on its own. It provided information on every email that was sent and all web activity. The downside to going through the logs was that it was a lot of data to look through. But there was a search function that allowed for you to look for keywords or sort the data to make it easier to find what you were looking for.

Cellebrite also provides a tab on any web content that it may find. In cases where it found something it provided the URL address and information on when the web page was accessed. In cases where no web content was provided it was usually due to having a basic device. Some of the mobile devices either were too basic to support web usage or contained web browser applications that were not too user-friendly.

3.4.2 Analysis using the T tool

With the T tool, which is similar to Autopsy (as mentioned before), the process for gathering email addresses and web usage information was not as user-friendly. There is a designated area where T places any email addresses that were found, but after some trial and error we figured out T contained a better method for finding email addresses. T has a tool that runs a search for an @ character and then places the results of that search into a file.

The way the search algorithm works is by looking for a pattern of some string of characters followed by an @ and then more characters followed by a final .com, .net, .gov, etc. We found that a lot of the output from this search resulted in text incorrectly identified as addresses, but many of those were obviously wrong and actual email addresses could be identified.

Web usage was tricky with the T tool. Similar to email content, there was an allocated area for T to place the results of web usage. We classified web usage as anything that suggested the device was used to connect to the Internet, such as stored bookmarks, cookies, or URLs. When web usage was not too apparent there was also a search method to be run where the algorithm searched for "www" followed by a URL pattern to try and find evidence of URLs.

3.5 Categorization

We categorized each phone based on the content and usage. This was a way to classify our findings and better understand different patterns found. We came up with seven different categories.

1. Very little to no content: phones that showed little or no content at all either because they were not used much or because content was successfully removed or deleted.
2. Normal user: phones that appeared to belong to a normal non-malicious user with the usual kinds of calls, messages, web usage, email, camera usage, etc.
3. Mostly Facebook: phones that mostly consisted of Facebook messages or Facebook content.
4. Basic Phone: seems like the phone belonged to a normal user, but the phone was too basic to have email or web usage.
5. High email activity: phones that showed a large use of email and not much else.

6. High web activity: phones that were mostly used for web and not much else.
7. Odd usage or content: phones whose logs represent non-normal usage, whose location seemed to change a lot, or contained odd content that did not obviously fit into any other category.

CHAPTER 4:
Results

4.1 Experimentation

For our experiment, we compared the analysis of mobile devices with Cellebrite versus T. We were looking for differences in content according to the output of both tools. We looked at all content in general, but focused on email addresses and web usage. We wanted to know if one tool reported more or less information on these specific types of files.

After gathering, the results from both tools were compared and the differences were measured.

4.2 Results

4.2.1 BZ-12 Samsung Galaxy S III

CPA reported 112 email conversations. Three conversations were found on the Gmail application from mail-noreply@google.com to mamourdu03@gmail.com, which belonged to a Micka' Mamour. The rest of the email conversations were found in the logs table and they were addressed to coupledelannee03@hotmail.fr which belonged to Mika Mik. Those emails were from various no-reply email addresses such as samsungaccount-noreply@samsung.com or billing@microsoft.com. There were also some emails that were gaming related such as those to xbox live, EA games, Black Ops 2, and Call of Duty. There was one Outlook account, the email content of which was mostly about gaming. All messages showed up as read. It looks like this phone was used for email from 8/18/2012 to 1/27/2013. When looking at the email content, we saw that most of the emails were confirmations for accounts for games.

Most of the web usage was connecting to a site to access a hotspot. Any other sites had .fr included in the address. There were also a few gaming blogs. Some bookmarks were ebay.com, facebook.com, google.com, nytimes.com, twitter.com, yahoo.com, fr.m.wikipedia.com, myspace.com, and www.weather.com. This phone had 655 calls logged, 861 SMS messages, over 40 contacts, over 6,000 images and 68 videos.

T reported that it found 542 email addresses using the script described in Chapter 3. Most of these matches were not actual email addresses, just matches to the keyword search script provided. A lot of them were vendor contact email addresses. T provides you with the amount of times a certain email came up in the keyword search. For example mamourd00@gmail.com came up the most at 36 times and then u0300@gmail.com came up secondmost at 18. After a closer look, it seems that there were only about 4 personal emails found.

The contacts seemed to be the same amount as CPA. T showed quite a bit more of deleted data than Cellebrite. The call log was significantly smaller at 27 and only about 4,000 images and 11 videos detected. We were not able to distinguish web usage. We classify this phone as one that belonged to a normal user. There was evidence of a significant amount of use to make phone calls and send SMS messages. There was also a large number of images reported by both CPA and T.

4.2.2 BZ-25 Samsung Galaxy Ace 3

CPA reported no email or web usage at all. Timestamps confirm that this phone was used from 2007-2008. Other data found was 1 user account, 28 SMS messages, 356 images, and 1 video.

T reported 152 emails. Only 2 seemed like actual email addresses, which were sinaidde-center4000@gmail.com which had 8 hits and ellenor1233@netlock.net with 3 hits. There was almost no evidence of web usage, but there were some Chromium cookies left behind which leads one to believe that the Chromium App was installed at some point. Other data it found was 269 images, and 2 videos. One would have to classify this phone as a basic phone, with the result that there was very little to no email or web usage.

4.2.3 CA-01 Apple iPhone

CPA reported one email address, andy1chiang1234@yahoo.com, which CPA identified as the user's AppleID. There were some cookies left from web usage which included google.com, twitter.com, wikipedia.com, and a lot from facebook.com. The phone had

88 contacts on Facebook and Facebook Messenger. All of the messaging was done on Facebook Messenger. There were over 8,000 pictures found, but most seemed to be system pictures. Other interesting data found was the location data, which all came from Virginia.

The T tool reported back that it found 0 email addresses but did find 6,196 matches to the keyword search. After a closer look it turns out none of those were actual personal email addresses, simply false matches to the keyword search. There was little evidence left of web usage. There were some cookies found. I was not able to see any of the Facebook data. The fact that there were no phone contacts and that they all came from Facebook makes me believe the user used this phone mostly for Facebook. There was some evidence of web usage but not much.

4.2.4 DE-18 Motorola Razor

CPA reported no evidence of web or email usage on this phone. All we were able to find were 70 SMS messages, 322 pictures, and 1 video. Timestamps suggest this phone was in use in 2006. The T tool produced an error message and was not able to analyze the contents of this phone. This phone was a basic phone. The lack of web or email use is most likely because of the fact that this phone is over 10 years old.

4.2.5 FR-04 Nokia Lumnia

This phone only provided a logical extraction. CPA found 6 personal pictures. Since a logical extraction does not provide a binary image, there was no image to be able to analyze with the T tool. We categorized this phone as having very little to no content.

4.2.6 FR-05 Apple iPhone

CPA reported no email addresses on this phone. The only thing we were able to see on this phone was that most of its location data suggested it was located in Europe. It also had 5 voicemail messages. I was not able to find any contacts or SMS messages.

The powering event data was really odd. The log suggests 8 powerups in the year 1970 and then jumps to one powerup in July of 2014, one in August 2014 and then 15 powerups in September 2014, of which 12 were within 2 hours of each other. The powerups shown for 1970 can be explained by the fact that 1970 is the default year for Unix-based systems. There were no applications installed on the device other than the default Apps.

T reported 3,127 email addresses, but those were only matches to the script. After further inspection, none were actual email addresses. Other than that, we were not able to get much from this phone. I would classify this phone as one with odd usage. The powerup data is not normal and the fact that there were no contacts or messages, or evidence of web usage, is odd. The phone was also named "phone repair" and it was linked to a PC named "PHONEREPAIR-PC," which suggests the phone might not been used as a traditional mobile phone.

4.2.7 IN-11 Dell ZTE Blade

CPA was able to detect one personal email address and there were cookies and stored bookmarks, which suggest web usage. The T tool displayed an error message and was not able to analyze the contents of this phone. It was classified as a phone with normal usage.

4.2.8 SG-27 Samsung Galaxy III

There were almost 200 email messages associated with the same single email address found by CPA. Most of the files found had been deleted. This phone was likely reset. There were 6 web bookmarks and 4 web cookies found suggesting web usage. The T tool reported an error when trying to import the binary files from this phone. It could not determine the file system type. It was classified as a phone with normal usage. There was a lot of other evidence that this phone was used normally and was reset, for example over 30,000 deleted SMS messages.

4.2.9 SG-28 LG Pop

This phone was imaged logically with CPA and it reported 475 SMS messages and 206 contacts. There was no email or web data reported. T was not able to provide an analysis since there were no binary files to import. It was classified as a phone with very little to no content.

4.2.10 SG-29 Nokia N97 Mini

CPA reported no email addresses and some web usage including 12 web cookies and 9 bookmarked sites. This phone was a Nokia with a Symbian OS and T was not able to analyze the binary file. It could not determine the file system type. It was classified as a phone with normal usage.

4.2.11 SG-34 Samsung Corby Pro

This phone was imaged logically. CPA found three pictures and nothing else. T was not able to provide an analysis since there were no binary files to import. It was classified as a phone with very little to no content.

4.2.12 SG-50 HTC Incredible S

CPA reported no email addresses, but a significant amount of web usage. There were over 30 sites bookmarked and almost 500 web cookies. A lot of files were deleted, which suggests the phone was reset. T got 4,500 hits with the keyword search, but only about 5 of those turned out to be legitimate personal email addresses. I would classify this phone as normal with high web activity.

4.2.13 SG-64 LG Optimus L3

CPA reported no email addresses or web usage. We did find saved evidence of connection to 34 wireless networks. Even though we did not find any URL addresses, the 34 saved networks could be a sign of web activity. A lot of the files looked like they were deleted, which suggests the phone might have been reset. T reported two personal email accounts found via the keyword search script and not much else. It was classified as a phone with normal usage.

4.2.14 SG-66 Nokia X3

CPA reported no email addresses. There were 6 web bookmarks and not much else. This phone was a Nokia and T was not able to analyze the binary file. It could not determine the file system type. It was classified as a phone with very little to no content.

4.2.15 SG-80 Apple iPhone

CPA was not able to find any email or web usage on this phone. It did recognize that it had a web browser application installed and some pictures but that is it. T found nothing but 84 matches to the keyword search; of those matches, most were email accounts but none seemed like personal ones. It was classified as a phone with very little to no content.

4.2.16 SG-81 Apple iPhone

CPA reported a specific email address as the user's Apple ID and 1 other email address associated with 30 inbox messages. There were 14 wireless networks, evidence of web history, and 169 web cookies found suggesting web usage was high on this phone. This phone was also heavily used for Facebook, as there were almost 500 Facebook contacts. T was able to find over 74,000 matches to the keyword search, but none seemed like legitimate personal email addresses. It was classified as a phone with high web and Facebook usage.

4.2.17 SG-88 Apple iPod

CPA found two Apple ID email addresses as well as 114 email conversations. This was the only device that was not a phone. There was a lot of evidence of web usage, there was some web history, web bookmarks, 5 IP connections, 4 wireless network records and over 4,000 web cookies. It was classified as a phone with high web usage.

4.2.18 TH-02 Sony Xperia

CPA reported no email addresses for this phone. It did find a lot of evidence of web usage. There were 19 wireless network records, 323 web cookies, 152 web bookmarks, and 309 web history entries. Classified under high web usage.

4.2.19 TH-05 BlackBerry Curve

CPA reported mostly a large call log on this phone. It found one email address, but it seemed to be a false positive. This was the first one that was not a valid personal email address. There was evidence of web usage such as 42 web history records and 5 web cookies. Also, 219 pictures and not much else. This phone was a BlackBerry and T was not able to analyze the binary file; it could not determine the file system type. It was classified as a phone with high web usage.

4.2.20 TH-09 Samsung Ch@t 322

All CPA found on this phone was 47 SMS messages that were deleted and nothing else. T was not able to find any useful data on this phone. It was classified as a phone with very little to no content.

4.2.21 TH-12 Apple iPhone

CPA reported no Apple ID unlike the other Apple devices. It did find over 500 email conversations all sent to one email address. Under user accounts it reported a SMTP and a POP service account both with the same user name as the email address. There was a lot of evidence of web usage, 334 web cookies, 29 web history, 20 network records, and 151 IP connections. T did identify 1 personal email address matching the one found with CPA in email conversations. It was classified as a phone with high web usage.

4.2.22 TH-20BlackBerry Curve

CPA reported no email activity and only 1 web bookmark. Other than that there were just a few pictures and 3 videos. This phone was a BlackBerry and T was not able to analyze the binary file. It could not determine the file system type. It was classified as a phone with very little to no content.

4.3 Categorization Results

The devices that were analyzed with CPA, and some with T as well, were placed in one of 7 categories described previously in Chapter 3. Table 4.1 shows the results as well as whether or not T was able to analyze a device. The devices were categorized based on the predominant usage of the devices reported from CPA and T.

Table 4.1. Categorization Results

Phone	Vendor	Name	Extraction Type	OS	T Extraction	Category
BZ-12	Samsung	Galaxy S III	Physical	Android	Y	Normal
BZ-25	Samsung	Galaxy Ace 3	Physical	Android	Y	Basic
CA-01	Apple	iPhone	Physical	iOS	Y	Facebook
DE-18	Motorola	Razor	Physical	Android	N	Basic
FR-04	Nokia	Lumnia	Logical	Windows	N	L/N content
FR-05	Apple	iPhone	Physical	iOS	Y	Odd
IN-11	Dell	ZTE Blade	Physical	Android	N	Normal
SG-27	Samsung	Galaxy III	Physical	Android	Y	Normal
SG-28	LG	Pop	Logical	Flash	N	L/N content
SG-29	Nokia	N97 mini	Physical	Symbian	N	Normal
SG-34	Samsung	Corby Pro	Logical	Proprietary	N	Normal
SG-50	HTC	Incredible S	Physical	Android	Y	Web
SG-64	LG	Optimus L3	Physical	Android	Y	Normal
SG-66	Nokia	X3	Physical	n/a	N	L/N content
SG-80	Apple	iPhone	Physical	iOS	Y	L/N content
SG-81	Apple	iPhone	Physical	iOS	Y	Facebook
SG-88	Apple	iPod	Physical	iOS	Y	Web/Email
TH-02	Sony	Xperia	Physical	Android	Y	Web
TH-05	BlackBerry	Curve	Physical	BlackBerry	N	Web
TH-09	Samsung	Ch@t 322	Physical	Android	Y	L/N content
TH-12	Apple	iPhone	Physical	iOS	Y	Web
TH-20	BlackBerry	Curve	Physical	BlackBerry	N	L/N content

Here, we show all the devices that were imaged and the category they were each placed in. We also show whether or not the devices were analyzed using T. "L/N content" means little to no content.

CHAPTER 5:
Conclusion and Future Work

5.1 Conclusion

We were able to extract a lot of data from multiple devices. We included a sample of those devices in this thesis. There were a few issues with the extraction process. A previous version of CPA was used due to the fact that an update on the hardware was not able to be installed. Some of the devices could not be imaged due to inability to charge, physical damage, or internal error. CPA did not provide physical extractions for some of the devices, so therefore we did a logical extraction. The devices that were imaged and analyzed allowed us to draw several conclusions: CPA and T can provide similar results for some devices, CPA had a better user interface, T was able to find more email addresses with its keyword search, T was only able to analyze images of Android and Apple devices, T could not analyze logically extracted devices, and web usage was easier to determine with CPA. But the tools used together could provide more data than one alone, and at least could provide confirmation of each other's results.

5.2 Future Work

We were only able to analyze a sample of the devices. Future work could include analysis of the rest of the devices and more. There were only devices from certain countries, and it would be good to include more countries. Also, analyzing the devices with updated versions of CPA's software might provide different results. We did not search the devices manually to try to verify results from either T or CPA. We did not analyze the devices with the Dirim system, so future work would include this as well.

THIS PAGE INTENTIONALLY LEFT BLANK

List of References

[1] A. Smith. (2015, Apr. 1). U.S. Smartphone use in 2015. Pew Research Center. Washington, DC. [Online]. Available: http://www.pewinternet.org/2015/04/01/us-smartphone-use-in-2015/. Accessed Aug. 9, 2016.

[2] Mobile phone forensics. (2016). Techopedia Incorporated. Edmonton, Alberta, Canada. [Online]. Available: https://www.techopedia.com/definition/2956/mobile-phone-forensics. Accessed Aug. 3, 2016.

[3] A. Smith. (2011, Aug. 15). Americans and their cell phones. Pew Research Center. Washington, DC. [Online]. Available: http://www.pewinternet.org/2011/08/15/americans-and-their-cell-phones/. Accessed Aug. 10, 2016.

[4] Mobile forensics: Smart phones. (2014). McCann Investigations. Houston, Texas. [Online]. Available: http://www.mccanninvestigations.com/mobile-forensics/. Accessed Jul. 1, 2016.

[5] Mobile forensics: Forensic tools. (2014, May. 3). National Institute of Standards and Technology. Gaithersburg, Maryland. [Online]. Available: http://csrc.nist.gov/groups/SNS/mobile_security/mobile_forensics.html. Accessed Aug. 3, 2016.

[6] R. Ayers, S. Brothers, and W. Jansen. (2014, May). Guidelines on mobile device forensics. *NIST Special Publication*. National Institute of Standards and Technology. Gaithersburg, Maryland. [Online]. Available: http://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-101r1.pdf. Accessed Jul. 20, 2016.

[7] V. Beal. (2011, Aug). Mobile operating systems (Mobile OS) explained. *Webopedia*. Quinstreet Enterprise. Foster City, California. [Online]. Available: http://www.webopedia.com/DidYouKnow/Hardware_Software/mobile-operating-systems-mobile-os-explained.html. Accessed Aug. 9, 2016.

[8] J. Drede. (2016, Jul. 4). What are the different types of Operating Systems? *wiseGeek*. Conjecture Corporation. Sparks, Nevada. [Online]. Available: http://www.wisegeek.org/what-are-the-different-types-of-mobile-phone-operating-systems.htm. Accessed Jul. 9, 2016.

[9] M. Yates, "Practical investigations of digital forensics tools for mobile devices," in *2010 Information Security Curriculum Development Conference*. ACM, 2010, pp. 156–162, accessed Aug. 4, 2016. Available: http://delivery.acm.org/10.1145/1950000/1940972/p156-yates.pdf?ip=205.155.65.226&id=1940972&acc=ACTIVE%20SERVICE&key=B318D1722F7F4203%2E44DF46464A4B769E%2E4D4702B0C3E38B35%2E4D4702B0C3E38B35&CFID=823837784&CFTOKEN=75091118&__acm__=1470808686_978ef8e5fd29b09394f0e51557e2980d

[10] Android (operating system). (n.d.). *Wikipedia*. [Online]. Available: https://en.wikipedia.org/wiki/Android_(operating_system)#Open-source_community. Accessed Apr. 28, 2016.

[11] A. Hoog, *Android forensics: investigation, analysis and mobile security for Google Android*. Waltham, Massachusetts: Elsevier, 2011.

[12] iPhone. (n.d.). *Wikipedia*. [Online]. Available: https://en.wikipedia.org/wiki/IPhone#Software. Accessed Apr. 28, 2016.

[13] BlackBerry OS. (n.d.). *Wikipedia*. [Online]. Available: https://en.wikipedia.org/wiki/BlackBerry_OS. Accessed Apr. 28, 2016.

[14] T. Storer, W. B. Glisson, and G. Grispos, "Investigating information recovered from re-sold mobile devices," in *Privacy and Usability Methods Pow-wow (PUMP) Workshop. ACM, University of Abertay, Dundee*, 2010, p. 2.

[15] Investigating Information Recovered from Re-sold Mobile Devices - Slides. (2010). University of Glasgow, School of Computing Science. [Online]. Available: http://scone.cs.st-andrews.ac.uk/pump2010/slides/storer.pdf. Accessed Nov. 16, 2016.

[16] M. Yates and H. Chi, "A framework for designing benchmarks of investigating digital forensics tools for mobile devices," in *Proceedings of the 49th Annual Southeast Regional Conference*. ACM, 2011, pp. 179–184.

[17] I. Yates *et al.*, "Practical investigations of digital forensics tools for mobile devices," in *2010 information security curriculum development conference*. ACM, 2010, pp. 156–162.

[18] iPhone analyzer. (2016). Slashdot Media. [Online]. Available: https://sourceforge.net/projects/iphoneanalyzer/. Accessed Aug. 30, 2016.

[19] Wordnik - Triage. (2016). Wordnik Company. [Online]. Available: https://www.wordnik.com/words/triage. Accessed Aug. 30, 2016.

[20] F. Marturana, G. Me, R. Berte, and S. Tacconi, “A quantitative approach to triaging in mobile forensics,” in *2011IEEE 10th International Conference on Trust, Security and Privacy in Computing and Communications*. IEEE, 2011, pp. 582–588.

[21] A. T. Ho and S. Li, *Handbook of digital forensics of multimedia data and devices*. West Sussex, United Kingdom: John Wiley & Sons, 2015.

[22] Physical extraction of mobile data. (2016). Cellebrite Corporation. Parsippany, New Jersey. [Online]. Available: http://www.cellebrite.com/Pages/physical-extraction-of-mobile-data. Accessed Sep. 10, 2016.

[23] Logical extraction of mobile data. (2016). Cellebrite Corporation. Parsippany, New Jersey. [Online]. Available: https://www.cellebrite.com/Pages/logical-extraction-of-mobile-data. Accessed Sep. 10, 2016.

[24] UFED Touch. (2016). Cellebrite Corporation. Parsippany, New Jersey. [Online]. Available: http://www.cellebrite.com/Mobile-Forensics/Products/UFED-Touch. Accessed Aug. 22, 2016.

[25] UFED Physical analyzer. (2016). Cellebrite Corporation. Parsippany, New Jersey. [Online]. Available: http://www.cellebrite.com/Mobile-Forensics/Applications/ufed-physical-analyzer. Accessed Aug. 22, 2016.

[26] S. L. Garfinkel, “Digital forensics xml and the dfxml toolset,” *Digital Investigation*, vol. 8, no. 3, pp. 161–174, 2012, accessed Aug. 22, 2016.

[27] S. L. Garfinkel. (2014). DFXML and other standards. [Online]. Available: https://calhoun.nps.edu/bitstream/handle/10945/44328/Garfinkel_2014-07-29%20DFXML.pdf?sequence=1&isAllowed=y. Accessed Aug. 22, 2016.

[28] Sleuthkit - Autopsy. (2003-2016). Sleuthkit Organization. [Online]. Available: http://www.sleuthkit.org/autopsy/. Accessed Aug. 21, 2016.

THIS PAGE INTENTIONALLY LEFT BLANK

Initial Distribution List

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California